Accelerating Performance Inference over Closed Systems by Asymptotic Methods

Giuliano Casale
Imperial College London
g.casale@imperial.ac.uk

ACM SIGMETRICS 2017
Closed queueing networks

\[ P(n) = \frac{1}{G} \prod_{i=1}^{K} n_i! \prod_{r=1}^{R} \frac{\theta_{ir}^{n_{ir}}}{n_{ir}!} \]

- \( K \) nodes
- \( R \) classes
- \( N \) jobs, \( N_r \) in class \( r \)

Product-form (\( \sigma_r = 0 \)):
Closed queueing networks

$K$ nodes
$R$ classes
$N$ jobs, $N_r$ in class $r$

Product-form ($\sigma_r = 0$):

$$G = \sum_\mathbf{n} \prod_{i=1}^{K} n_i! \prod_{r=1}^{R} \frac{\theta_{ir}^{n_{ir}}}{n_{ir}!}$$

See paper for case $\sigma_r > 0$. 
Motivating example

- Problem: service demand inference
  - $Q_{ir}$: empirical mean queue-length

- Maximum-likelihood estimation (MLE)

$$\max_{\theta \in \Omega} \sum_{i,r} Q_{ir} \log \theta_{ir} - \log G$$

where $\theta \equiv (\theta_{ir})$ and $G \equiv G(\theta)$. 

$\theta_{ir}$ service demand, $N$: jobs, $R$: classes, $K$: queues
Related work

Computing the normalizing constant \( G \):

- **CA**: Convolution algorithm
- **RECAL**: Recursion by chain
- **MoM**: Method of moments
- **RAY**: Ray method
- **MCI**: Monte Carlo integration
- **TE**: Taylor expansion
- . . .
Motivating example

- 60s run, 3 nodes, 3 classes, 120 jobs

This prompts us to revisit $G$’s computational theory.
Consider $f(x)$ and data points $x = \theta_1, \ldots, \theta_K$

Newton’s interpolation polynomial

$$N(x) = \prod_{i=1}^{K} (x - \theta_i) \cdot \left[ \theta_1, \ldots, \theta_K \right] f(x) + \cdots$$

Explicit form for divided differences

$$[\theta_1, \ldots, \theta_K] f(x) = \sum_{k=1}^{K} \frac{f(\theta_k)}{\prod_{i\neq k}(\theta_k - \theta_i)}$$
Single-class models

- Explicit form for single-class models \((R = 1)\)

\[
G = \sum_n \prod_{k=1}^{K} \theta_{nk} = \sum_{k=1}^{K} \frac{\theta_k^{N+K-1}}{\prod_{i \neq k}(\theta_k - \theta_i)}
\]

- Matching the definitions

\[
G = [\theta_1, \ldots, \theta_K]x^{N+K-1}
\]

\(\theta_{ir}\) service demand, \(N\): jobs, \(R\): classes, \(K\): queues
A new integral form

- Hermite-Genocchi formula for divided differences

\[ G = \frac{(N + K - 1)!}{N!} \int_{\Delta_K} (\theta_1 u_1 + \ldots + \theta_K u_K)^N \, du \]

on the simplex \( \Delta_K = \{|u| = 1, u \geq 0\} \).

- Laplace transform on volume of \( \Delta_K \) yields the McKenna-Mitra integral form for \( G \) over \( \mathbb{R}_+^K \).
Example

\( K = 2 \) queues
\( u_1 + u_2 = 1 \)
\( \theta_1 > \theta_2 \)

\[ G \propto \int_{\Delta_K} (\theta_1 u_1 + \theta_2 u_2)^N \, du \]

\( \theta_{ir} \) service demand, \( N \): jobs, \( R \): classes, \( K \): queues
Multiclass analysis

- We generalize the result to multiclass models

\[ G = \frac{(N + K - 1)!}{N_1! \cdots N_R!} \int_{\Delta_K} \prod_{r=1}^{R} (\theta_{1r} u_1 + \ldots + \theta_{Kr} u_K)^{N_r} \, d\mathbf{u} \]

- Proof uses multinomial theorem and

\[ P(n) \propto n_1! \cdots n_K! \propto \int_{\Delta_K} \prod_{k=1}^{K} u_k^{n_k} \, d\mathbf{u} \]

\( \theta_{ir} \) service demand, \( N \): jobs, \( R \): classes, \( K \): queues
Computational methods

From the integral form, we derive:

- Explicit solutions
- Asymptotic expansion
- Cubature rule
Computational methods

From the integral form, we derive:

- Explicit solutions
- Asymptotic expansion
- Cubature rule
Explicit solutions

- Integral form for single class models ($R = 1$)

$$G = \frac{(N + K - 1)!}{N!} \int_{\Delta_K} (\theta_1 u_1 + \ldots + \theta_K u_K)^N \, d\mathbf{u}$$

- Integral form for multiclass models ($R > 1$)

$$G = \frac{(N + K - 1)!}{N_1! \cdots N_R!} \int_{\Delta_K} \prod_{r=1}^{R} (\theta_{1r} u_1 + \ldots + \theta_{Kr} u_K)^{N_r} \, d\mathbf{u}$$

$\theta_{ir}$ service demand, $N$: jobs, $R$: classes, $K$: queues
From products to sum

- Can we relate the two cases?

\[
\prod_{r=1}^{R} x_r^{N_r} = \frac{1}{N!} \frac{\partial^{N_1} \ldots \partial^{N_R}}{\partial t_1^{N_1} \ldots \partial t_R^{N_R}} \left( \sum_r t_r x_r \right)^N
\]

where \( N = \sum_r N_r \).

- Multiclass integrand reduced to single class!
$O(N^R)$ Explicit form

- Replace derivatives with finite differences.
- Applying to the integrands

$$G = \sum_{t:0 \leq t_r \leq N_r} \prod_{r=1}^{R} \left( \frac{(-1)^{N_r-t_r}}{N_r!} \right) \begin{pmatrix} N_r \\ t_r \end{pmatrix} \sum_{k=1}^{K} \prod_{i \neq k} \frac{\theta_k^{N+M-1}}{(\theta_i - \theta_k)}$$

where $\theta_k = \sum_r t_r \theta_{kr}$.

- Same time complexity as CA, but $O(1)$ space
$O(N^K+1)$ Explicit form

- Applying the idea to products in $P(n)$

$$G = \sum_{h \geq 0 : h \leq N} \frac{(-1)^{N-h}}{N_1! \cdots N_R!} \binom{N + K - 1}{N - h} \prod_{r=1}^{R} \left( \sum_{k=1}^{K} h_k \theta_{kr} \right)^{N_r}$$

- Same time as RECAL, but $O(1)$ space

$\theta_{ir}$ service demand, $N$: jobs, $R$: classes, $K$: queues
From the integral form, we derive:

- Explicit solutions
- Asymptotic expansion
- Cubature rule
Logistic transformation

\[ u_1 \rightarrow \frac{e^x}{1 + e^x} \]

\[ u_2 \rightarrow \frac{1}{1 + e^x} \]

\[ \left| \frac{\partial u}{\partial x} \right| = \frac{e^x}{(1 + e^x)^2} \]

- Approximate mapping to a Gaussian over $\mathbb{R}^{K-1}$
- Close to Gaussian also for small $N$
Properties

- What are the parameters of the Gaussian?
- **Mean** $\hat{u}$ obtained iteratively

$$\hat{u}_i = \sum_{r=1}^{R} \frac{N_r}{N} \sum_k \theta_{kr} \hat{u}_k$$

- Asymptotically identical to **mean-value analysis**
- Explicit form found for the **covariance matrix** $A$

$\theta_{ir}$ service demand, $N$: jobs, $R$: classes, $K$: queues
Heavy-load asymptotics

- We force $\epsilon N$ self-looping jobs at each node.
- For small $\epsilon$, we get in heavy load ($N \to \infty$)

$$G \sim \frac{(N + K - 1)!}{N_1! \cdots N_R!} \sqrt{\frac{(2\pi)^{K-1}}{\det(A)}} \prod_{r=1}^{R} \left( \sum_{k=1}^{K} \theta_{kr} \hat{u}_k \right)^{N_r} \prod_{i=1}^{K} \hat{u}_i$$

- We refer to this result as logistic expansion (LE).

$\theta_{ir}$ service demand, $N$: jobs, $R$: classes, $K$: queues
Example: MLE revisited

- 60s run, 3 nodes, 3 classes, 120 jobs

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean Abs. Perc. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>40</td>
</tr>
<tr>
<td>MCI</td>
<td>118</td>
</tr>
<tr>
<td>MoM</td>
<td>71</td>
</tr>
<tr>
<td>RAY</td>
<td>59</td>
</tr>
<tr>
<td>REC</td>
<td>77</td>
</tr>
<tr>
<td>TE</td>
<td>95</td>
</tr>
<tr>
<td>NoG</td>
<td>96</td>
</tr>
<tr>
<td>LE</td>
<td>1</td>
</tr>
</tbody>
</table>
Computational methods

From the integral form, we derive:

- Explicit solutions
- Importance sampling
- Asymptotic expansion
- Cubature rule
Cubature rules

\[ \int_{\Delta_K} f(u) \, du \approx \sum_j w_j f(u(j)) \]
Grundmann-Möller cubature

- $O(S^K)$ samples symmetric to simplex barycenter

$$u^{(j)} = \left(\frac{(2b_1^{(j)} + 1)}{(2S + K - 2s)}, \ldots, \frac{(2b_K^{(j)} + 1)}{(2S + K - 2s)}\right)$$

$\forall s = 0, \ldots, S$ and $b^{(j)} \geq 0$ s.t. $|b^{(j)}| = S - s$.

- For $G$, the rule is exact when $S = \lceil (N - 1)/2 \rceil$

- Lower degree rules give approximations of $G$
Grundmann-Möller cubature

degree $S = 1$

4 points

$u_1$

$u_2$

$u_3$

error $= 92.9\%$
Grundmann-Möller cubature

Degree $S = 2$

18 points

$\text{error} = 61.0\%$
Grundmann-Möller cubature

degree $S = 4$

35 points

error $= 6.0\%$
Grundmann-Möller cubature

degree $S = 6$

84 points

error = 0.0%
Validation: Light-load

- $K, R \in [2, 32], \frac{N}{R} = 2$, 1125 MLE problems

Cubature rule best in light load.
Validation: Heavy-load

- $K, R \in [2, 32], \frac{N}{R} = 40, 1125$ MLE problems

Logistic expansion best in heavy load.
Conclusion

Main findings:

- Novel integral form for $G$ on the unit simplex
- Explicit solutions
- Asymptotic expansions and cubature rule

Additional results in the paper:

- Importance sampling method
- Models with delay nodes ($\sigma_r > 0$)
- Validation on random models